Indian J. Phys. **73B** (2), 191-201 (1999)

IJP B
— an international journal

Locating critical points and identifying fragmentation channels in noble gas clusters by genetic algorithm

Pinaki Chaudhury and S P Bhattacharyya

Department of Physical Chemistry, Indian Association for the Cultivation of Science, Jadavpur, Calcutta-700 032, India

Received

Abstract In this paper, we discuss the global optimization technique of Genetic Algorithm, both its operational aspect and the basic theorem which guides its evolution We also discuss the strong and weak features of the algorithm and mention strategies which can circumvent the weak features. We apply the technique to search out global minimum energy structures of neutral Xe clusters, mixed Xe and Ar clusters and try to predict probable modes of fragmentation in an ionic cluster Ar_{30}^{+}

Keywords Genetic algorithm, optimization techniques, Noble gas clusters.

PACS Nos. 36 40 -c, 32.10.-f

1. Introduction

Global optimization is a subject of interest to many workers. The search for novel and efficient global optimization techniques are pursued by many researchers. Genetic Algorithm (GA) [1-3] is one such truly global optimization technique. The most commonly used search techniques are usually deterministic and heavily depend on the starting point of search. The success of such methods in finding the global geometry in most cases is fortuitous. Hence, one needs to understand and use methods which employ stochastic search procedures and can escape being trapped in any local minima. These methods must also be truly unbiased so that they are able to arrive at the global geometry starting from initial configurations which are generated randomly and are structurally quite different from the global minimum energy structure.

Many global methods are currently in use. The techniques of Simulated annealing [4,5], Basin Hopping Monte Carlo [6], Molecular dynamics on deformed Potential Energy Surfaces [7] and Genetic Algorithm have been used extensively and have performed impressively.

The specific problem of finding global geometry in neutral and ionic rare gas clusters have attracted the attention of many workers [8-13]. These systems are known to possess innumerable probable configurations. For example, a cluster of moderate size N = 13 is known to possess 988 minima. The ionic clusters for which a lot of experimental data are available, also

possess multiple minima, but shows the appearance of magic numbers *i.e.* particular configurations with extra stability. In this paper, we propose to use the technique of GA in solving global minimum search problem in some neutral clusters (Xenon), mixed Xenon-Argon clusters and an ionic cluster (Ar_{30}^{+}) which undergoes fragmentation and tries to achieve extra stability by generating structures that conform to magic numbers.

2. The method

Genetic Algorithms employ a search procedure which borrows concepts from natural selection and natural genetics. It starts with a population of probable solutions called strings which undergo evolution by application of certain operators to produce a newer set. Certain selection criteria are employed which pick up the better or fitter strings and discard the rest. This evolution is continued for a number of steps, called generation in GA terminology, till the fitness of the strings reach a maximum value or the solution to the problem is found out. In most GA applications the strings are binary coded, but we have used a real coded GA in the structural optimizations we have undertaken.

For an n atom cluster there are $3n$ coordinates $(q_1, \ldots, q_{3n})$ to be optimized. We generate N randomly generated potential solution vectors $s_1, s_1, \ldots s_n$.

Typically the *k-th* vector has the following structure

$$s_k = (q_1^k, q_2^k, \ldots, q_m^k, \ldots, q_{3n}^k)^t. \tag{1}$$

The potential of interaction of the *k-th* cluster is

$$V_k = \sum_{i<j=1}^{n} \sum_{j=1}^{n} V(i, j), \tag{2}$$

where $V(i, j)$ is the Lennard Jones potential between two neutral Ar atoms when both i and j represent neutral argon atoms and $V(i, j)$ represents the $Ar^+ - Ar$ potential when one of the two indices (i, j) refer to an Ar^+ ion. The fitness f_k of the k-th string is computed by taking

$$f_k = e^{-\sqrt{\sigma_k}} \tag{3}$$

where

$$\sigma_k = (V_k - V_L)^2 \tag{4}$$

V_L is a lower bound to the energy and can be upgraded, if needed.

Why a GA is able to find out an optimal solution can be understood by understanding how the fundamental GA operators (i) selection (ii) crossover and (iii) mutation work.

(i) Selection :

The role of the selection operator is to scoop out relatively fitter solutions from the existing set and create a new pool with increased average fitness. Many selection procedures are reported in contemporary literature. Some of the common ones are Roulette Wheel based selection, Boltzmann selection [14, 15], Rank selection [16] and Tournament selection [17]. We have employed the roulette wheel based procedure. The central idea is to allocate a slot in a roulette

wheel to each of the N strings proportional to its fitness. Next we call a random number N times. That string is selected for which the random number lies in the respective slot. As the slot width is proportional to the fitness of a string, after selection we get only the fitter ones (multiple copies of a particular string may be produced). The fitter pool is now ready for crossover.

(ii) Crossover :

From the pool a certain percentage of the strings (guided by crossover probability, p_c) are selected for crossover. After selection the strings are grouped in pairs, randomly. Next one exchanges information between the two strings. If the m^{th} site of the k^{th} string is chosen as the crossover site then the strings s_k and s_l where

$$s_k = q_1^k, q_2^k, \ldots\ldots, q_m^k, \ldots\ldots, q_n^k$$

$$s_l = q_1^l, q_2^l, \ldots\ldots, q_m^l, \ldots\ldots, q_n^l, \tag{5}$$

become

$$s_k' = q_1^k, q_2^k, \ldots\ldots fq_m^k + (1-f)$$

$$s_l' = q_1^l, q_2^l, \ldots\ldots fq_m^l + (1-f)q_m^k, \ldots\ldots, f\,q_n^l + (1-f)q_n^k. \tag{6}$$

Generally 50% to 75% of the strings are subjected to crossover and the rest unchanged.

(iii) Mutation :

After crossover the strings are subjected to mutation. Mutation is a unary operation and this occurs with a very low probability $p_m = 0.001$.

Under it the vector

$$s_k' = q_1^k, q_2^k, \ldots\ldots q_m^k, \ldots\ldots, q_n^k \tag{7}$$

becomes

$$s_k'' = q_1^k, q_2^k, \ldots\ldots q_m^{k''}, \ldots\ldots, q_n^k \tag{8}$$

where

$$q_m^{k''} = q_m^k + (-1)^L \Delta. \tag{9}$$

The action of these three operators on a population ends one generation. Why these three operators transform a rather poor and randomly generated solution set to the actual solution may not be easy to guess off hand. There is a theorem called Schema Theorem which explains this evolution [3]

The Schema theorem :

In our formulation of the problem, we have used a real coded GA. However, to introduce the Schema Theorem, it would be prudent to use binary representation. Suppose we have a string

$$A = 0111000.$$

Now we define an useful concept, the concept of a schema. A schema is a similarity template describing a subset of strings with similarities at certain string positions. Let us define a

schema $H = *11*0**$ where * refers to a don't care symbol or wild symbol which can match either a 0 or a 1 at a particular position. String A is an example of schema H as the string positions match at the fixed points 2, 3 and 5.

We also define two properties of a schema, (i) order and (ii) defining length. Suppose we have two schema.

$$Sc_1 = 011*1**$$

and

$$Sc_2 = 0******$$

The order O is simply the number of fixed positions in a string and the defining length δ is the distance between the first and last fixed string positions.

Hence $\delta(Sc_1) = 4$ and $\delta(Sc_2) = 0$

and

$O(Sc_1) = 4$ and $O(Sc_2) = 1$.

If $m(H, t)$ is the number of strings in the t-*th* generation having the particular schema (H) then the number in the $(t+1)$-*th* generation can be shown to be [3]

$$m(H,t+1) \geq m(H,t)\frac{f(H)}{\bar{f}}\left[1 - p_c\frac{\delta(H)}{l-1} - O(H)p_m\right], \tag{10}$$

where

$f(H)$ = fitness of a string containing schema H,

$\bar{f}$ = average fitness of the population,

p_c = probability of crossover,

(H) = defining length of a schema,

l = length of the string,

$O(H)$ = order of a schema,

p_m = probability of mutation.

Interpretation of the Schema growth equation :

It is evident from the schema growth equation that a schema with higher than average fitness, low order and small defining length will survive and grow in number as the population evolves.

3. Results and discussions

We have carried out GA optimization of neutral clusters (Xe_{10}, Xe_{13}, Xe_{15} and Xe_{20}), a xenon cluster doped with an argon atom $Xe_{12}Ar$ and we have followed the optimization of an ionic cluster Ar^+_{30}, which fragments into a Ar^+_{29} and an Ar atom to gain stability.

(i) Neutral Clusters Xe_n (n = 10, 13, 15, 20)

The interaction potential among the Xe atoms is given by the standard Lennard Jones potential

$$V = 4\varepsilon\sum_{i,ji<j}\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}, \tag{11}$$

where $r_{ij}'s$ are the distances between the Xe atoms and σ and ε are standard LJ parameters for Xe – Xe interaction ($\sigma_{Xe-Xe} = 4.10$, $\varepsilon_{Xe-Xe} = 1.8525$). Figure 1(a) shows the optimization profile for Xe_{15} with number of generations. Figures 1(b), (c), (d) and (e) depict the structures and energies of Xe_{10}, Xe_{13}, Xe_{15} and Xe_{20}.

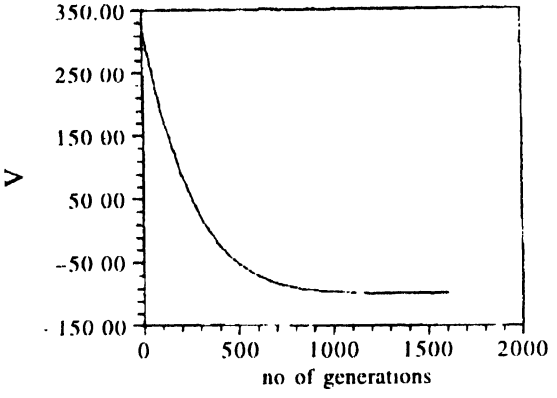


Figure 1(a) The variation of potential V corresponding to the best individual string with the number of generations during the search for the global minimum structure of Xe_{15}.

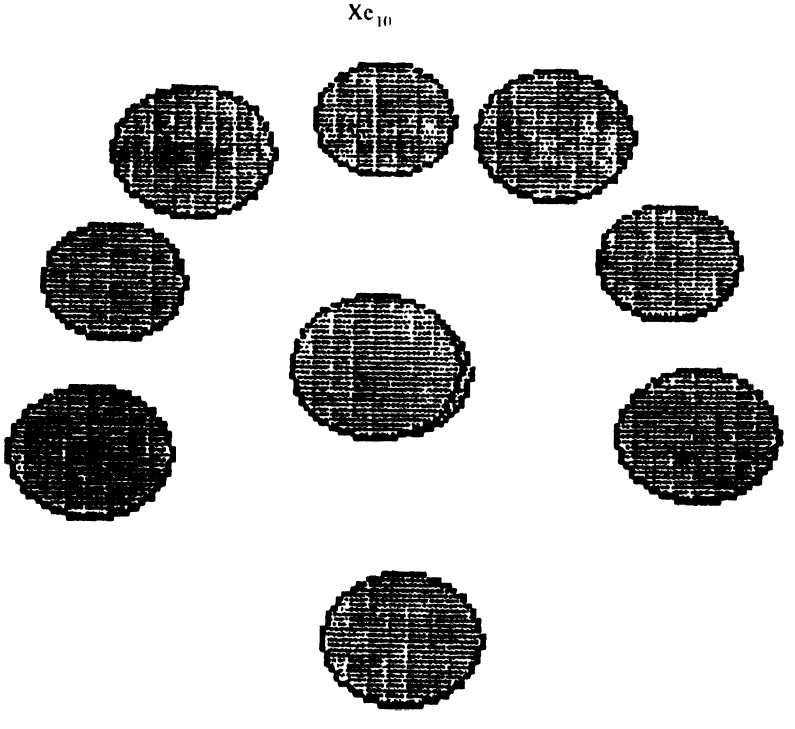


Figure 1(b). The global minimum structure of Xe_{10} cluster.

(ii) Mixed Xe – Ar clusters :

For optimization of clusters of the type Xe_nAr_m the only modification to be made in the potential is to include Xe – Ar interaction. The hetero interactions can also be modelled by Lennard

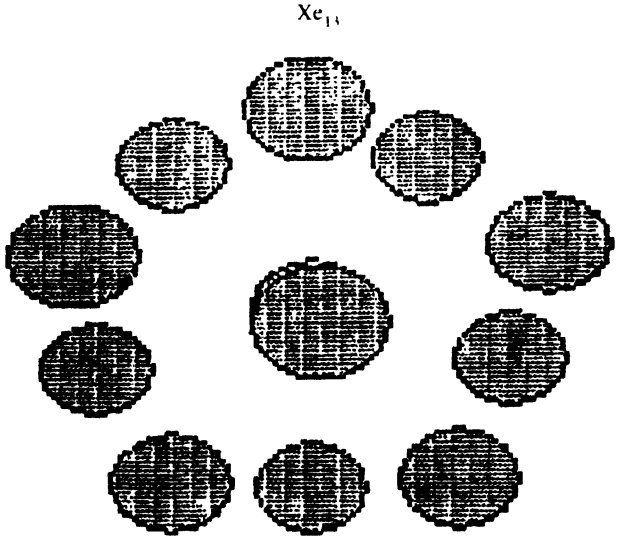


Figure 1(c). The global minimum structure of Xe_{13} cluster

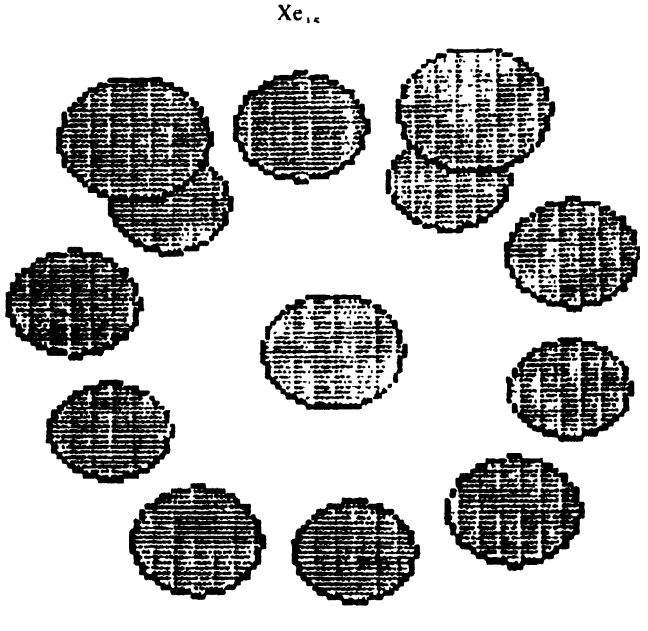


Figure 1(d). The global minimum structure of Xe_{14} cluster

Jones potential with appropriate ε and σ ($\sigma_{Xe-Ar} = 3.65$, $\varepsilon_{Xe-Ar} = 1.48$). We have optimized a heterocluster of the type Xe_{12} Ar. The structure and energy of the cluster are depicted in Figure 2.

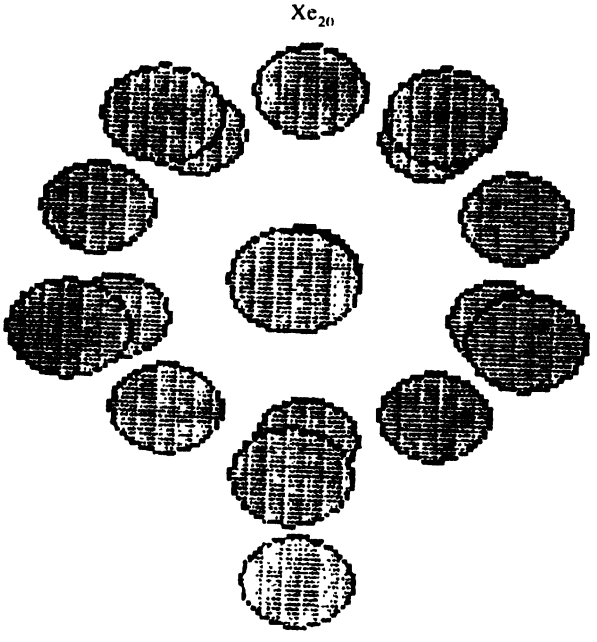


Figure 1(e). The global minimum structure of Xe_{20} cluster

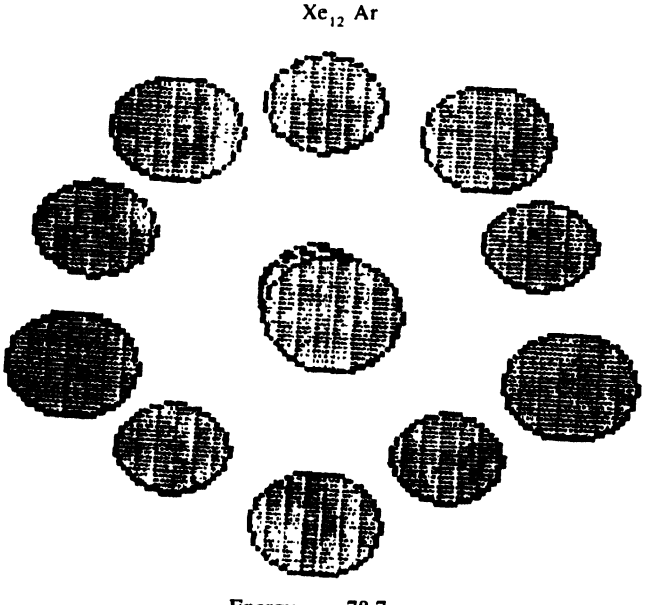


Figure 2. The global minimum structure of $Xe_{12}Ar$ cluster.

(iii) Ionic Cluster Ar_{30}^{+} :

To optimize Ar_n^+ type of clusters we need to find out the potential energy curve for Ar_2^+. This is done by ab-initio quantum chemical methods. Ar_2^+ has been treated as a 15 valence electron problem [18, 19]. Relativistic effective core potential has been used to represent the effects of the core. For the valence electrons $4s4p1d$ basis set has been used for the SCF calculations at different internuclear distance. SCF calculations have been done on Ar_2 and an extensive CI has been done with energy selected configurations at each internuclear distance of the ion. The ground state potential has fitted to a polynomial in $\frac{1}{R}$ [20]. For Ar – Ar interactions, the standard Lennard Jones Potential is used.

Ar_{30}^{+}, during the course of optimization is seen to be rather unstable and dissociates into an ionic core and a neutral atom. This instability of Ar_{30}^{+} has been reported in many experimental findings [21]. Figure 3(a) shows the geometry along with the energy of the fully optimized structure while Figures 3(b) and 3(c) show two intermediate structures which seem to retain all its atoms. So Ar_{30}^{+} which is one unit in the beginning ejects an argon atom to gain extra stability by achieving a magic number configuration as the structure evolves following ionization.

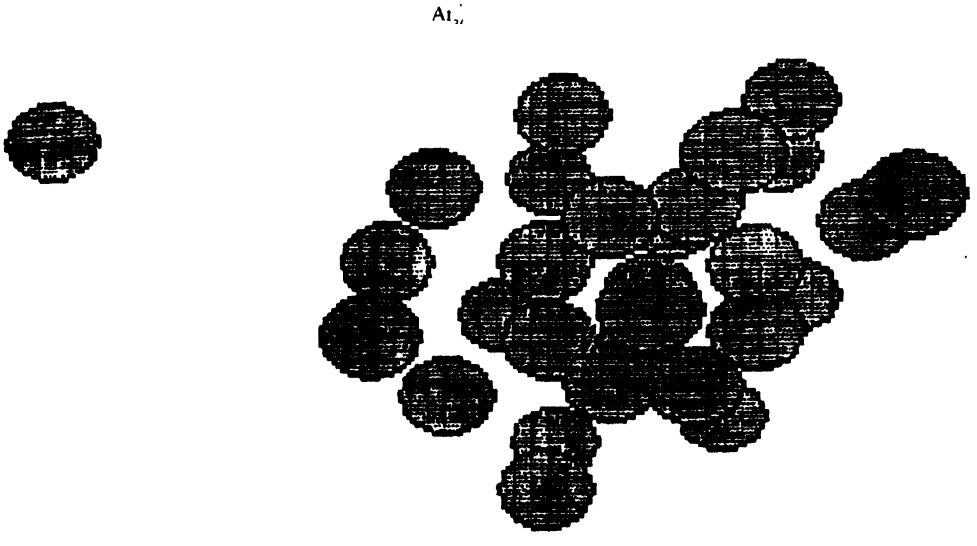
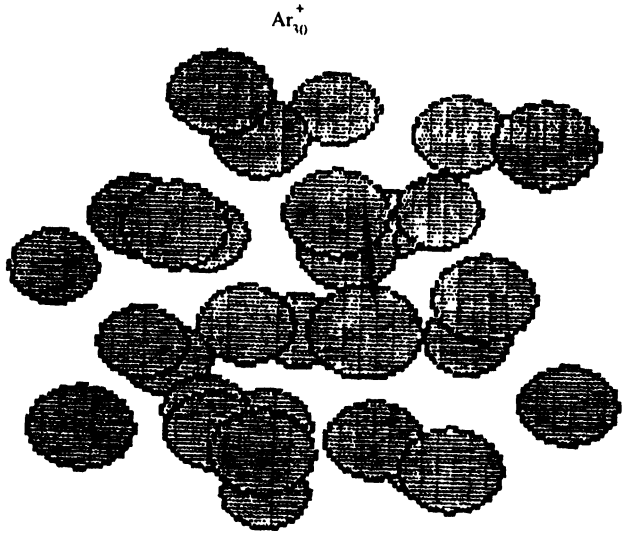


Figure 3(a). The global minimum structure of fragmented Ar_{30}^{+} towards the end of the evolution. Note that fragmentation has taken place.

4. Strong and weak features of the algorithm

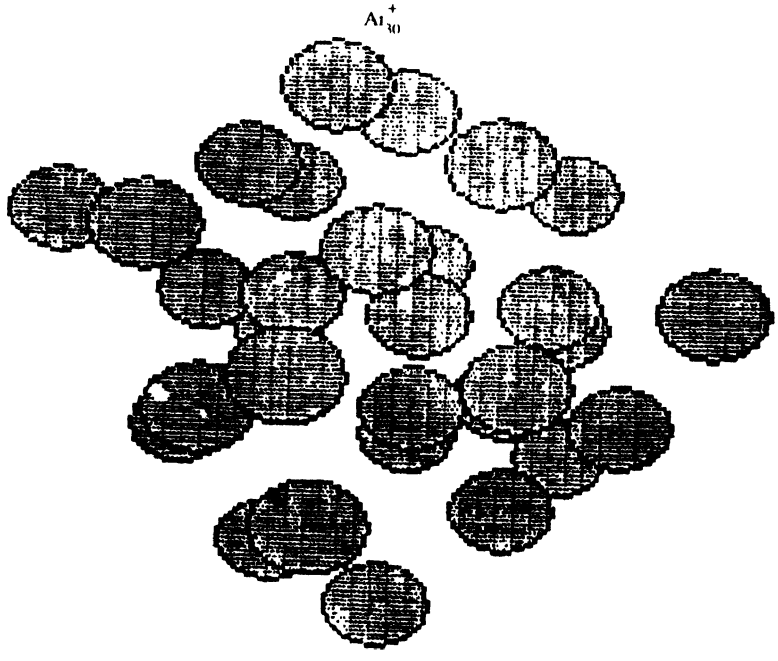
First we examine the strong features of the algorithm. GA can be an effective tool if the space to be searched is large, is known not to be perfectly smooth and unimodal, or if the fitness function is noisy. However one can argue that other stochastic optimizers can also be used, say for example Simulated Annealing. However GA has the special advantage of starting with a large trial solution set, which already has the potential of exploring a large volume of the

search space. Moreover the operators, crossover and mutation enables the GA to search better points within the search space and to expand the domain of search to include unexplored regions. So one can say that GA simultaneously exploits existing information and explores new regions. The other available methods all start with one probable solution, improves it in



Energy = - 564 96 Kj/mol

Figure 3(b) The structure of Ar_{30}^{+} during the course of evolution after 100 generations



Energy = - 820 51 Kj/mol

Figure 3(c). The structure of Ar_{30}^{+} during the course of evolution after 300 generations

successive trials. Reported results of calculations on Lennard Jones clusters with simulated annealing are not impressive and global structures have been found for small systems only.

The above discussion seems to suggest that GA is a potent technique and one might be tempted to use it in its simplest version. However GAs suffer from operational difficulties, the principal one being premature convergence. During evolution, sometimes the solution set converges to a local minima and cannot evolve further as the diversity in the solution set is lost. This results primarily due to enhanced selection pressure. Simple roulette wheel selection procedure is not advisable in these situations and the use of Boltzmann type selection as used in Simulated Annealing might be helpful. Another way might be to design a diversification operator, which is similar to mutation in operation, which perturbs the solution set at periodic intervals and help maintain diversity. Our experience shows that if we can maintain diversity in the population and give sufficient time for the set to evolve, the search reaches the global solution.

5. Conclusions

We have shown that a Genetic Algorithm, which is rather simple in its design and operation can be a powerful tool in optimizing various kinds of noble gas clusters of intermediate size. It has the potential of predicting possible modes of fragmentation of an ionic cluster. The algorithm might require certain modifications, which are problem specific, but on the whole it is a powerful technique and worth further investigation [22].

Acknowledgments

One of us (P.C.) wishes to thank the Council of Scientific & Industrial Research, Government of India, New Delhi for the award of a Senior Research Fellowship.

References

[1] D E Goldberg *Genetic Algorithms in Search, Optimisation and Machine Learning* (Reading, MA Addison-wesley) (1989)

[2] F Hoffmeister and T Back *Genetic Algorithms and Evolution Strategies, in Proceedings of the First International Conference on Parallel Problem Solving from Nature (PPSN)* Dartmund, Germany eds H P Schwefel and R Manner, (1990)

[3] John H Holland *Adaptation in Natural and Artificial Systems* (MIT, USA : The MIT Press) (1998)

[4] S Kirkpatrick, C D Gelatt and M P Vecchi *Science* **220** 671 (1983)

[5] P Dutta and S P Bhattacharyya *Phys. Lett* **A148** 331 (1990)

[6] D J Wales and J P K Doye *J. Phys. Chem.* **A101** 5111 (1997)

[7] J Pillardy and L Piela *J. Phys Chem* **99** 11805 (1995)

[8] S K Gregurick, M H Alexander and B Hartke *J. Chem. Phys.* **104** 2684 (1996)

[9] R S Judson, E P Jaeger, A M Treasurywala and M L Peterson *J. Comp. Chem.* **14** 1407 (1993)

[10] D M Deaven, N Tit, J R Morris and K M Ho *Phys. Rev. Lett* **75** 288 (1995)

[11] D M Deaven, N Tit, J R Morris and K M Ho *Chem. Phys. Lett.* **256** 195 (1996)

[12] J A Northby *J. Chem. Phys.* **87** 6166 (1987)

[13] W J Pullan *J. Comp. Chem.* **18** 1096 (1997)

[14] M de la Maza and B Tidor *Boltzmann Weighted Selection Improves Performance of Genetic Algorithm* (A I Memo 1345, MIT Artificial Intelligence Laboratory) (1993)

[15] A Prugel-Bennett and J L Shapiro *Phys. Rev. Lett.* **72** 1305 (1994)

[16] J E Baker *Proceedings of the Fourth International Conference on Genetic Algorithms and their Applications* Erlbaum ed. J J Grefenstette (1985)

[17] D E Goldberg and K Deb *Foundations of Genetic Algorithm.* eds G Rawlins and Morgan Kaufmann (1991)

[18] L F Pacios and P A Christiansen *J Chem Phys* **82** 2664 (1985)

[19] P C Hariharan and J A Pople *Chem. Phys Lett* **66** 217 (1972)

[20] P Chaudhury, K K Das and S P Bhattacharyya (to be published)

[21] J C Phillip *Chem. Rev* **86** 619 (1986)

[22] P Chaudhury and S P Bhattacharyya *Chem Phys* (in press)